

www.medkharbach.com

ASSESSMENT LITERACY

Simply Explained for Teachers



M E D K H A R B A C H , P H D

Introduction

In this short guide, I explain what assessment literacy is and why it matters for teachers and educators. The idea for this guide took shape while I was reading James Popham's book *Assessment Literacy for Educators in a Hurry* (2018). In that book, Popham makes a bold claim: a large number of educators are not assessment literate because they lack an understanding of the fundamental concepts and procedures of educational assessment. I agree with him to a large extent.

Most of us did not receive enough training in the science of assessment and testing. Teacher education programs focused mainly on pedagogy and instructional practice. We studied teaching methods, classroom management, and pedagogical theories, but assessment often received only brief attention. Many of us designed our first classroom tests with little guidance. Experience helped, and practice led to improvement, but assessment literacy as a specialized body of knowledge was largely missing from our preparation.

I decided to address assessment literacy for two main reasons. The first is the gap Popham points to. Many teachers lack this form of literacy, and strengthening it is long overdue. The second reason has to do with AI, especially generative AI, which complicates assessment in fundamental ways.

With powerful language and content generation tools now widely available, no assessment can truly claim to be AI-proof. Students can easily rely on generative AI for many types of assignments. This reality forces us to revisit core questions about assessment: what exactly are we assessing, and how should we assess it? I argue that developing strong assessment literacy can help educators respond to these challenges with more clarity and confidence.

The purpose of this guide is to offer a concise introduction to assessment literacy. I draw mainly on the work of two established scholars in the field: James Popham (2018) and Richard Stiggins (1995). I begin by discussing Popham's definition of assessment literacy. I then introduce the key concepts that make up assessment literacy, explore the qualities of assessment-literate educators, and conclude with a discussion of assessment quality standards.

1. What is Assessment Literacy?

When we talk about literacy, we usually default to reading and writing. In educational work, though, literacy carries a broader meaning. It points to the ability to develop competence and knowledge in a specific area. As Popham (2018) puts it, literacy refers to developing "competence and knowledge in a specific arena" (p. 4).

Seen through this lens, assessment literacy refers to a particular kind of professional knowledge. It is concerned with how educators understand assessment and how that understanding shapes their instructional and evaluative decisions. Popham describes assessment literacy as “an individual’s understanding of the fundamental assessment concepts and procedures deemed likely to influence educational decisions” (p. 2).

Two elements in this definition deserve careful attention: concepts and procedures.

Assessment concepts form the conceptual backbone of assessment literacy. These include ideas educators encounter repeatedly in practice and policy discussions, such as validity, reliability, fairness, and formative assessment. Each of these concepts helps teachers judge the quality of assessment information and interpret what assessment results actually say about student learning.

Assessment procedures, on the other hand, refer to how assessments are built, administered, and evaluated. Popham explains that procedures “refer to the techniques or methods commonly used to build or evaluate tests” (p. 4). This includes decisions about item formats, scoring approaches, feedback practices, and the conditions under which assessments take place.

In the sections that follow, we look more closely at some of the foundational assessment concepts and explore how they shape everyday assessment practice.

1.1. Validity

Validity is one of the most frequently used terms in assessment literature, yet it is also one of the most misunderstood. Popham defines validity as “the degree to which evidence and theory support the interpretations of scores for proposed uses of tests” (p. 17). This definition immediately shifts attention away from scores themselves and toward how those scores are interpreted and used. Validity asks a practical question educators face every day: do the conclusions we draw from assessment results make sense, given the evidence we have?

Popham leaves little room for treating validity as a technical side issue. He writes that “without validity, educational testing would have no point, no purpose, and no legitimate application. In this chapter, we will explore why” (p. 17). In classroom terms, an assessment that lacks validity cannot support sound instructional decisions, no matter how carefully it is scored or how neatly results are reported.

One helpful way Popham frames validity is through the idea of making the invisible visible. He notes that “validity is the bedrock of our confidence in testing’s ability to make the covert overt” (p. 18). What students have learned and internalized remains covert. It lives in their thinking, understanding, and reasoning. Assessment attempts to make that learning overt by eliciting evidence we can observe, such as responses, performances, or products. Validity concerns how well that evidence actually reflects the learning we care about.

1.2. Reliability

Reliability often enters assessment conversations through a simple idea: consistency. Popham makes this connection explicit when he explains that “the degree to

which a test is reliable is the degree to which that test measures with consistency” (Popham, 2018, p. 35). When educators talk about reliability, they are asking a straightforward question. If the same learning stays the same, does the assessment yield similar results?

Popham draws an important distinction between reliability and validity. He notes that “unlike validity, assessment reliability resides not in a test-based inference but in the test itself. Accordingly, educators should expect that the tests they use will be accompanied by evidence regarding reliability” (p. 36). Reliability focuses on the stability of measurement, not on the meaning we assign to scores. It asks how dependable the assessment results are before any interpretation takes place.

This idea becomes clearer in Popham’s concise definition of reliability as “the consistency with which a test measures whatever it measures” (p. 38). The phrase “whatever it measures” is key here. A test can measure something consistently and still fail to measure what educators actually care about. Reliability supports measurement stability, but it does not guarantee meaningful conclusions on its own.

1.3. Fairness

Fairness brings assessment literacy into direct contact with students’ lived experiences. It asks a practical and ethical question at the same time: does this assessment give all students a genuine opportunity to show what they know and can do?

Popham (2018) explains that evidence of fairness in educational tests comes from two main sources. The first is judgmental evidence of test fairness. Here, educators and reviewers examine assessment items closely and ask whether any of them introduce bias. As Popham puts it, “if a test is composed of items that are biased, then the test itself will be biased. That is, it will not measure fairly” (p. 57).

The second source is empirical evidence of fairness. This approach looks directly at student performance data across subgroups. Popham describes it this way: “this approach boils down to a per-item contrast in the success rates of different subgroups of students. If, for instance, Item 32 on a grade-to-grade promotion test in reading was answered correctly by 86 percent of 3rd grade boys but only 61 percent of 3rd grade girls, then we ought to take a gander at Item 32 to see if it contains content that might give 3rd grade boys an advantage over 3rd grade girls. If so, it will need a rewrite” (p. 60). Patterns like these do not automatically prove bias, but they signal the need for closer inspection.

Fairness, as Popham argues, is also supported through three closely related concepts that guide assessment design and use: accessibility, universal design, and accommodation.

Accessibility focuses on removing unnecessary barriers so students can show what they know. As Popham explains, it “refers to the need to ensure test-takers have an unobstructed opportunity to demonstrate their prowess regarding whatever is being assessed” (p. 63). Universal design addresses accessibility at the planning stage. Popham describes it as building assessments “in a way that will maximize accessibility for all students” (p. 64). The aim is to anticipate learner variability from the start so fewer fixes are needed later.

Accommodation comes into play when adjustments remain necessary. Popham defines accommodations as “adjustments made at any point in the measurement process to provide students with a fair opportunity to respond to a test without altering the fundamental nature of what’s being assessed” (p. 64). He illustrates this with a clear example. Increasing font size for a visually impaired student on a reading test supports access without changing what the assessment aims to measure. The accommodation preserves the core construct while removing an avoidable barrier.

1.4. Formative Assessment

Popham (2018) defines formative assessment, or assessment for learning, as “a planned instructional process in which assessment-collected evidence is used either by teachers or by students to make any necessary adjustments in their current efforts” (p. 93).

That definition does a lot of work. When you slow down and look at its components, you start to see why formative assessment is often misunderstood in classrooms.

First, formative assessment is planned. It does not just happen on the spur of the moment. Popham is very clear on this point. If it is not planned in advance, then it does not qualify as formative assessment in the proper sense. Second, formative assessment is not a test. It is a process. Tests may be part of that process, but they are simply tools for collecting evidence. As Popham puts it, it is “a serious labeling mistake to refer to a test as ‘a formative assessment’” (p. 89).

Third, formative assessment is ongoing. It unfolds across instruction, feeding information back into teaching and learning while there is still time to adjust. The whole point is to inform next steps, not to certify learning after the fact.

Key Takeaways

- Assessment literacy refers to a specific form of professional knowledge that shapes how educators design, interpret, and use assessments.
- It rests on two core elements: understanding key assessment concepts and knowing how assessment procedures work in practice.
- Validity focuses on the quality of the interpretations we make from assessment results, not on the test itself.
- Reliability concerns consistency. An assessment should produce stable results before any interpretation takes place.
- Fairness asks whether all students have a genuine opportunity to show what they know, free from unnecessary barriers or bias.
- Accessibility, universal design, and accommodation support fairness by removing obstacles without changing what is being assessed.
- Formative assessment is a planned, ongoing process that uses evidence of learning to adjust teaching and learning in real time.
- Assessment literacy helps educators make better decisions, draw more defensible conclusions, and support student learning more effectively.

2. Characteristics of Assessment Literate Teachers

Assessment-literate educators approach assessment with clarity. As Stiggins (1995) explains, they come to any assessment knowing “what they are assessing, why they are doing it, and how best to assess the learning they care about” (p. 241). They understand how to gather sound evidence of student performance and can anticipate what might go wrong. When problems arise, they know how to prevent them before they affect results.

Assessment-literate educators also understand the role of formative assessment in learning. Popham (2018) notes that they “know the value and use of formative assessment” and make better decisions about which assessments to use to understand students’ learning progress and how to act on the results (p. 11). Assessment information guides instruction, feedback, and support.

Another key quality is judgment. Stiggins (1995) emphasizes that assessment-literate educators know the difference between sound and unsound assessments. They can recognize when an assessment provides trustworthy information and when it does not.

Finally, assessment-literate educators understand quality standards (Stiggins, 1995). They know how to meet specific standards of assessment quality and use those standards to guide how assessments are designed, selected, and interpreted. Validity, reliability, and fairness serve as practical reference points in their everyday work.

3. On the Importance of Assessment Literacy

As I mentioned before, assessment is key for both teachers and students. Being assessment literate allows you to design tests that measure what should be measured. Assessment literacy also enables you to cultivate the knowledge necessary to make sound instructional decisions.

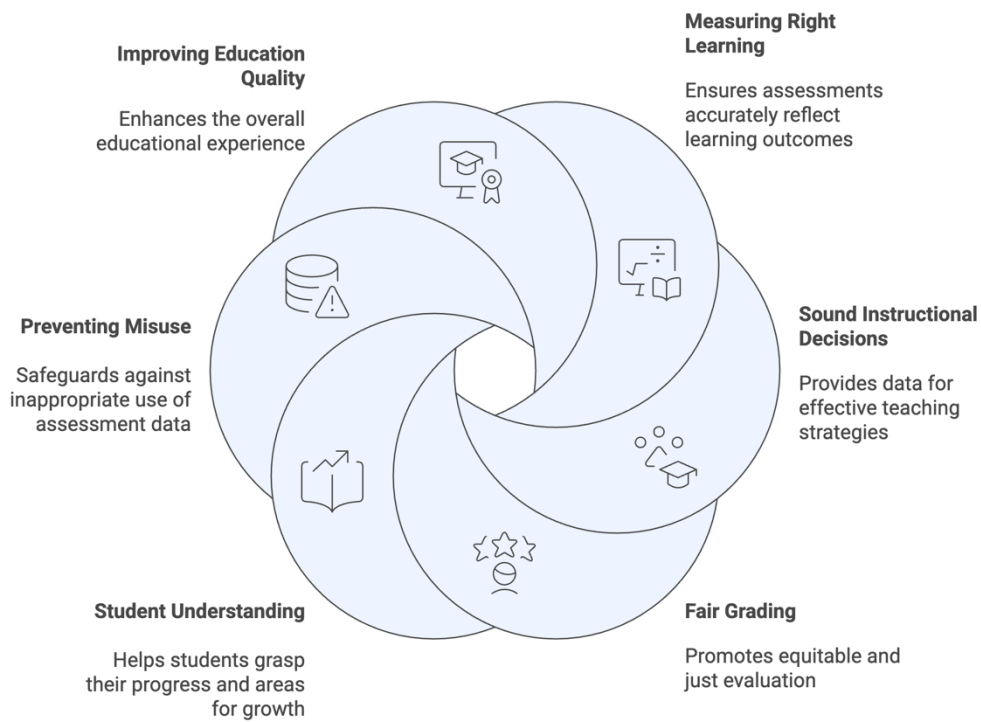
For students, having an assessment-literate teacher enhances their learning and ensures they are treated fairly in educational measurement. Scores, grades, and teacher judgments can truly reflect their achievement and help them learn about their learning, including their strengths and weaknesses.

Additionally, being assessment literate helps you avoid three main mistakes that, according to Popham (2018, p. 7), assessment-illiterate educators make:

1. **Using the wrong tests.** Tests that do not genuinely detect students' learning progress can communicate erroneous messages about the effectiveness of teaching, or lack of it.
2. **Misusing results of the right tests.** This involves misapplication, that is, applying tests designed for a specific purpose to a different purpose.
3. **Failing to employ instructionally useful tests.** Some assessments generate data but provide little actionable guidance for instruction.

As Popham (2018, p. 11) states: "Assessment literacy, in short, improves the likelihood that the students we serve will receive a better education. And this, of course, is an aspiration well worthy of our pursuit."

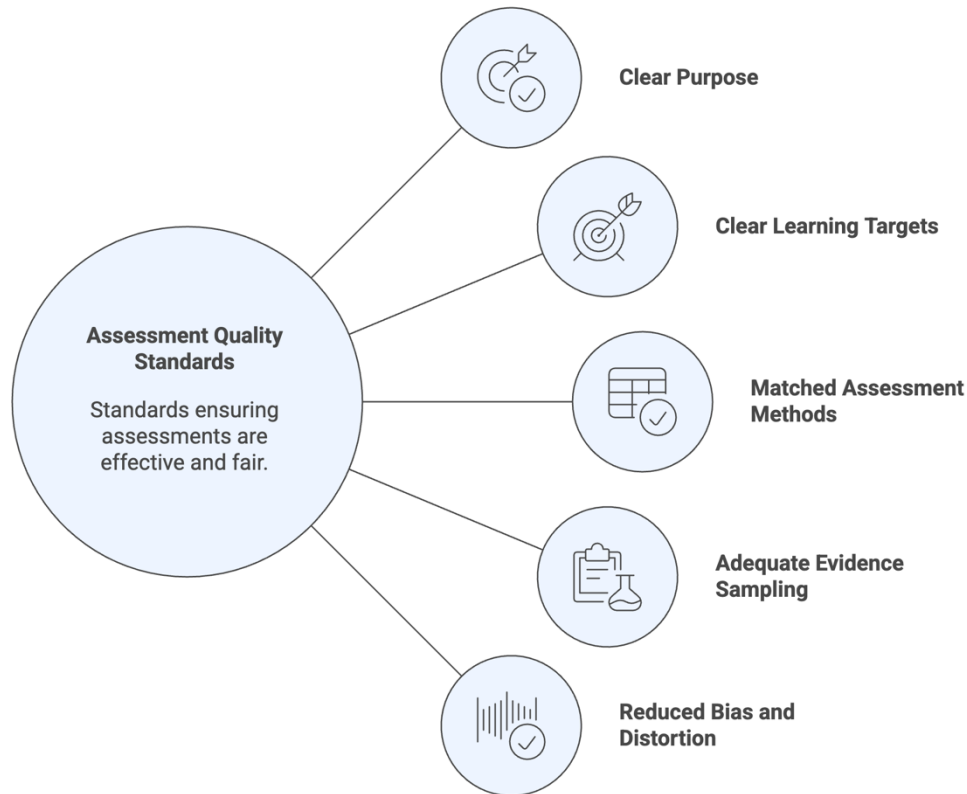
The Importance of Assessment Literacy



4. Standards of Assessments Quality

Stiggins (1995) offers a practical way to think about assessment quality by laying out five core standards that guide sound assessment practice. These standards reflect everyday decisions teachers make when designing, choosing, and using assessments in real classrooms. They also provide a framework for thinking carefully about why we assess, what we assess, how we assess, and how confident we can be in the conclusions we draw. The sections that follow walk through each of these standards and connect them directly to classroom practice.

Assessment Quality Standards (Stiggins, 1995)



1. Starting with clear purposes

Think about all the different reasons you assess students throughout a week. You might be checking whether your lesson clicked, figuring out which students need more support, deciding how to group kids for tomorrow's activity, or gathering evidence for report cards. You also share assessment information with parents, colleagues, and administrators, each of whom uses it differently.

Stiggins' key insight is that no single assessment can serve all these purposes well. That quick exit ticket you use to plan tomorrow's mini-lesson? It's perfect for that moment, but it won't give you what you need for a parent conference or a grade-level planning meeting. The data your school collects for improvement plans operates at a completely different level than the close-up view you need to help an individual student move forward.

Strong assessment practice starts with being intentional. Before you design or choose an assessment, get clear on who will use the information and what decisions it will inform.

Classroom-level work demands detailed, nuanced insight into each learner. School or district-level decisions need broader patterns across many students.

2. Focusing on achievement targets

Stiggins' second emphasis focuses on being clear about the achievement targets we want students to reach, and this has very practical implications for classroom assessment. He reminds us that achievement is not one thing. Students are expected to develop content knowledge, learn how to reason with that knowledge, perform skills, and, at times, create high-quality products. Knowledge still matters deeply. Skills and problem-solving rest on it. A student cannot communicate in another language, conduct a science investigation, or write a strong argument without a solid base of vocabulary, concepts, and procedures.

At the same time, Stiggins pushes teachers to rethink what it means to “know” something. Mastery is not limited to memorizing facts. In today's world, knowing how to locate, evaluate, and use information matters just as much. That insight opens the door to different assessment approaches, including open-book tasks and take-home work, when the goal is to see how students use knowledge and not recite it. Still, learning content only for its own sake misses the point. We want students to use what they know to reason, solve problems, and make decisions.

3. Selecting proper assessment methods

Stiggins' third emphasis is choosing assessment methods that match the learning target. Because students are expected to develop different kinds of achievement, no single assessment method can capture everything. Assessment-literate teachers draw from multiple options, including selected-response tasks, short or extended writing, performance assessments, and personal communication such as questioning, conferencing, or discussion.

The key is alignment. Knowledge targets are often best assessed with selected-response items, essays, or conversations, while skills and performances require students to actually do the task so teachers can observe proficiency. Reasoning may show up across several formats, but only when the task fits the type of thinking being assessed. Stiggins' core message to teachers is simple and practical: clarify the learning target first, then choose the assessment method that gives the clearest evidence of that target, rather than relying on a single format or following trends.

4. Sampling student achievement

Stiggins' fourth emphasis highlights the importance of sampling student learning carefully. Any assessment captures only a small slice of what students know or can do, so quality depends on choosing tasks that fairly represent the learning target. The size and type of the sample should match the assessment method, the breadth of the target, and the stakes of the decision.

Narrow targets require fewer tasks, while broader goals and high-stakes decisions call for larger, more representative samples. Stiggins' reminder to teachers is that confidence in assessment results comes from thoughtful sampling, not from a single question or task standing in for everything students have learned.

5. Avoiding bias and distortion

Stiggins' final emphasis focuses on avoiding bias and distortion in assessment. Even well-designed assessments can misrepresent learning when factors such as language proficiency, reading and writing demands, emotional state, health, or testing conditions interfere with students' performance.

Problems can also come from the assessment itself, including unclear questions, weak prompts, or inconsistent scoring. These issues can lead teachers to draw inaccurate conclusions about what students actually know or can do. Ultimately, assessment literacy includes anticipating these sources of bias and taking steps to reduce them, so assessment results reflect student learning as accurately and fairly as possible.

What Is Assessment Literacy?

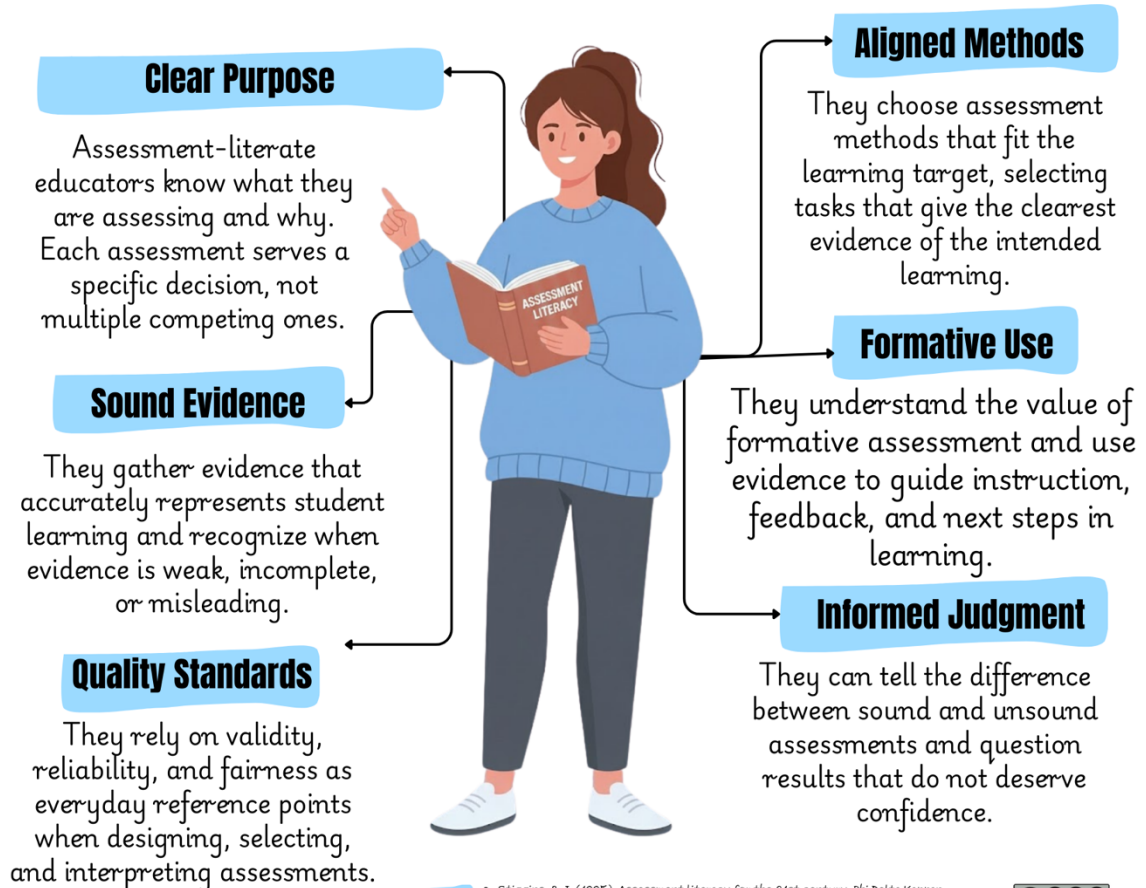


According to Popham (2018), assessment literacy “consists of an individual’s understanding of the fundamental assessment concepts and procedures deemed likely to influence educational decisions.” (p. 2)

Prepared by Med Kharbach, PhD

www.educatorstechnology.com

Qualities of Assessment Literate Teachers



Sources

- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238-245. <https://www.jstor.org/stable/20402338>
- Popham, W. J. (2018). Assessment literacy for educators in a hurry. ASCD.



Conclusion

Assessment literacy is a core part of our professional practice as teachers and educators. As this guide has shown, assessment literacy brings together clear thinking about purpose, a solid grasp of foundational concepts, careful attention to fairness and quality, and informed judgment about evidence of learning.

Drawing on the work of Popham and Stiggins, this guide has framed assessment literacy as a form of professional knowledge that shapes everyday classroom decisions. It influences how assessments are designed, how results are interpreted, and how information is used to support learning.

In a time when assessment practices are under pressure from accountability demands and new technologies such as generative AI, assessment literacy matters more than ever. It gives educators a way to respond thoughtfully rather than reactively. It helps teachers ask better questions about what they assess, how they assess it, and why it matters. Ultimately, assessment literacy supports better teaching and more meaningful learning. It increases the likelihood that assessment information reflects students' actual achievement and helps both teachers and students move learning forward. That goal alone makes assessment literacy a pursuit worth sustained attention.

References

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238–245. <https://www.jstor.org/stable/20405538>

Popham, W. J. (2018). *Assessment literacy for educators in a hurry*. ASCD.